

ЗБОРНИК

1

ДИЈАГНОСТИЧКИ МОДЕЛИ

Увод

Показателот на нивото на современиот развиток на природните науки е даден со степенот на примената на математички достигнувања, кибернетиката и современата компјутерска техника.

Во тој поглед медицината мошне заостанува зад техничките науки, иако во последно време примената на методите и резултатите на прецизните науки оди со забрзано темпо.

Интересно е дека и покрај древната старост и на медицината и на математиката, взаемниот интерес на една за друга, се јавува од неодамна со појавата на електронските компјутерски машини и примената на кибернетиката во медицината.

До половината на овој век имало само неколку работи од оваа област, а со наглиот подем на електронската компјутерска техника и нивниот број забрзано се зголемува. Јасно е дека тоа е последица на необичната сложеност на објектот што се изучува — човекот, а исто така последица на специфичноста на клиничката медицина, која преставува комбинација на науката и практиката.

Користењето на математичките методи и електронските сметачки машини (во понатамошниот текст ЕСМ), во дијагностиката на болестите дава значаен придонес во развојот на современата медицина.

Благодарејќи на кибернетичките методи, некои медицински гледишта можат да се третираат на следниот начин:

Организмот претставува многу сложен систем за себе, кој е составен од колосален број на различни атоми и молекули обединети во ќелии, ткива, органи и функционални системи. Познато е дека меѓу елементите доаѓа до размена на енергија, а со тоа и до промена на структурата. Заправо таа размена на енергијата и промена на структурата во време од одредени зависности е функција на организмот.

Состојбата на организмот е определена со карактеристика која се наоѓа во самата структура на системот. Јасно е дека има многу состојби, вклучувајќи ги оние состојби кои се јавуваат како последица на дејството на надворешниот свет врз организмот.

Спрема тоа, болестите се всушност подмножество на множеството состојби, кога доаѓа до промена на функцијата на организмот.

Сега да го разгледаме односот: болниот — лекар.

После историјатот и физичкиот преглед на болниот, најчесто се потребни натамошни лабораториски тестови (тоа се на пример, тест на крвта, тест на функцијата на дробовите, тест на количество на шеќер во крвта и други).

Веќе на овој чекор имаме помош на ЕСМ. Врз основа на историјатот на болниот и неговата моментална состојба, ЕСМ може да го определи минималниот комплет на тестови, кој е потребен да се изврши, што многу му помага на лекарот за утврдување на вистинската дијагноза, а исто така и на пациентот ослободувајќи го од непотребни тестирања.

Јасно е дека при одредувањето на овој комплет на тестови мора да се води сметка да се отфрлат само оние тестови кои навистина не се поврзани со конкретниот случај, односно, веројатноста на одредувањето на вистинската дијагноза да не се најде под даденото ниво.

Пред да се употребат ефективно ЕСМ во аспектот на медицинската дијагноза, мора, покрај другото, да биде познат и мисловниот процес на еден лекар за поставување на дијагнозата.

Ако прашаме еден лекар: „Како доаѓате до медицинската дијагноза?“ — неговото објаснување би било слично на следното:

— прво, добивам податоци во врска со случајот од историјатот на болниот, потоа од физичкиот преглед и од лабораториските тестови,

— второ, ги групирам релативно важните и различните симптоми и знаци — некои се примарни, а некои помалку важни,

— трето, за да дадам точна дијагноза, морам да имам во главата комплетна листа на дијагнози кои се специфични за случајот, така да добивам една комплетна претстава. Тогаш избирам една дијагноза, на друга, се додека не ја добивам онаа која одговара на дадениот случај, или некоја нејзина варијанта, или во друг случај, точната природа не може да се определи.

Јасно е дека ова е груб, упростен процес на поставување на дијагноза. Ова е, всушност, само скица на тој процес.

Спрема тоа, општата задача на медицинската дијагностика лежи во определувањето на карактерот на болеста, врз основа на посматраните симптоми, резултатите на испитувањето и историјатот на болниот.

Зи примена на ЕСМ при поставување на медицинската дијагноза, неопходно е информациите врз основа на кои лекарот поставува дијагноза, да бидат запишани во одреден облик. Исто така, неопходен е и математички запис на самиот процес на поставување на дијагнозата.

Примената на математиката во дијагностиката, пред се, се состои во тоа да со помош на ЕСМ или со специјални дијагностички табели врз основа на определен избор на симптоми, да се одреди дијагноза која има најголем можен степен на веројатност.

Од теоријата на веројатноста познато е дека веројатноста за појава на некој настан е број ограничен помеѓу нула и единица, односно, процентуално меѓу 0 и 100%.

Според обемни истражувања, утврдено е дека дијагнозата добиена по горниот пат е прифатлива за медицината, ако нејзината веројатност е поголема од 80%.

Нека $S = \{S_i / i \in I\}$ е множество на симптоми, S_i i -ти симптом, I — индексно множество, а аналогно на тоа $D = \{D_j / j \in I\}$ е множество на дијагнози.

Ако сите симптоми кои можат да се сретнат кај посматраниот болен ги распоредиме по определен редослед, можно е да се даде математички запис на состојбата на организмот. Во симптомите ги вклучуваме и историјатот на организмот и лабораториските резултати и податоците добиени со физичкиот преглед.

Со S_k ќе ја означиме појавата на k -тиот симптом, а со S_k — отсуството на k -тиот симптом. Тогаш резултатите на испитувањето можат да се претстават во облик на n — димензионалниот вектор:

$$(S_1 \bar{S}_2 \bar{S}_3 \dots S_i \bar{S}_i + 1 \dots \bar{S}_n)$$

(претпоставка е дека имаме n — симптоми).

Врз основа на медицинските знаења, на овој вектор може да се придружи некој елемент од множеството D , значи некоја дијагноза D_j :

$$D_j \sim (S_1 \bar{S}_2 \bar{S}_3 \dots S_i \bar{S}_i + 1 \dots \bar{S}_n)$$

Горниот запис само ја регистрира присутноста или отсуството на симптомите. Ако внимаваме на редоследот на симптомите со воведувањето на ознаките:

$S_i = 1$ — ако i -тиот симптом се јавува,

$\bar{S}_i = 0$ — ако i -тиот симптом не се јавува,

(за секое $i \in I$), добиваме релација еквивалентна на претходната:

$$D \sim (1 \ 0 \ 0 \dots 1 \ 0 \dots 0).$$

Јасно е дека и овој запис само ја регистрира присутноста или отсуството на симптомите, сите симптоми се со иста важност и интензитет.

ОПШТО ЗА МОДЕЛИТЕ

Алгоритам е множество од правила кои ја одредуваат содржината и редоследот на операциите за обработка на дадените податоци, до бараниот резултат.

Ако алгоритмот се користи за распознавање на болестите на организмот, тогаш се вика дијагностички алгоритам. Понатаму, во изложената материја се дадени неколку приоди кон проблемот на изнаоѓање на дијагнозите, неколку дијагностички алгоритми — модели.

Логичкиот модел ги користи знаењата на математичката логика; алгоритмот на веројатноста, правилата, законите и резултатите од теоријата на веројатноста; методот на фазниот интервал по идеја е близок на логичкиот модел, додека статистичкиот метод се состои во статистичко споредување на комплексот на симптомите (на синдромот), откриени кај болниот во случаите проучени порано и забележани во клиничка архива.

Геометриската интерпретација е обид проблемот да се илустрира во Еуклидовиот n — димензионален простор.

Секој од овие модели има свои предности и недостатоци. Можно е да се комбинираат повеќе модели, посебно кога постои сомневање во точноста на дијагнозата, а исто така, кога по еден алгоритам добиеме дека повеќе дијагнози го задоволуваат критериумот.

I. МОДЕЛ НА ФАЗЕН ИНТЕРВАЛ

1. За моделот

Во основата на методот лежи претпоставката дека состојбата на организмот и неговата функција можат да се опишат со помош на определен систем од параметри: X_1, X_2, \dots, X_n . Во тој случај на здрав организам одговараат едни вредности на параметарот, а на болен организам други вредности.

При претпоставка тие параметри да се оски на некој координатен систем, имаме дека множеството на вредности се јавува како точка од тој систем. Значи, секоја состојба на организмот е можно да се претстави со точка од тој простор.

Да ја означиме со A областа на здрава, нормална состојба на организмот. Аналогно, точките кои претставуваат определени болни состојби, формираат некои други области B_1, B_2, \dots

На тој начин, исти заболувања кај различни луѓе се класифицираат како блиски, но сепак тоа се различни состојби.

Овие области имаат променлива густина на веројатноста. Нека со $P(X_1, X_2, \dots, X_n)$ ја означиме густината на веројатноста така да $pdx_1, pdx_2, \dots, pdx_n$ се определени веројатности кога при дадената патолошка состојба параметрите на системот се наоѓаат помеѓу x_1 и $x_1 + dx_1, x_2$ и $x_2 + dx_2, \dots, x_n$ и $x_n + dx_n$.

Како мерка на дадена патолошка област, може да се земе збир на дисперзиите на отстапувањата од центарот (т.е. од математичките очекувања) по параметрите x_1, x_2, \dots, x_n .

$$D_B^2 = \sum_{i=1}^n D^2 x_{iB} = \sum_{i=1}^n (x_i - x_B)^2$$

каде што x_B е математичко очекување на областа B .

Ако претпоставиме дека состојбата на некој од организмите не е во потполност определена со апсолутно значење на параметрите $x_i' = x_i - x_{i0}$, па во првата апроксимација на независни случајни величини ќе имаме:

$$D_B^2(x_i') = D_B^2(x_i) + D^2 x_{i0} = D_B^2(x_i) + D_A^2$$

бидејќи параметрите $x_{10}, x_{20}, \dots, x_{n0}$ ја определуваат нормалната состојба на организмот, односно имаме:

$$D_B^2(X_i) = D_B^2(X_i') - D_A^2$$

Од оваа формула следува дека ако во својство на координати на овој простор земеме неапсолутно значење на параметрите x_1, x_2, \dots, x_n и нивното отстапување од индивидуалната норма, тогаш областа на нормалната состојба се стеснува во точка, а мерката на патолошката област се намалува за мерка на областа на нормалната состојба, односно и таа област се стеснува.

Ако посматрањето на болниот почнува кога неговата состојба се наоѓа во точката M_1 , тогаш треба да се определи на која област и припаѓа таа точка, т.е. кажано со јазикот на медицината треба да се определи дијагнозата.

Обично, припадноста на точката M_1 е позната за толку координати (x_1, x_2, \dots, x_n) , колку се јавуваат при испитувањето на болниот. Со тоа што останатите координати се непознати и положбата на точката M_1 е неопределена.

Всушност, тука имаме множество од точки со геометриско место $M_1(x_1, x_2, \dots, x_n)$ т.е. имаме не точка, туку површина M_1 .

Ако таа површина сече точно една патолошка област, тогаш без оглед на отсуство на параметрите за определување на останатите координати на точката M_1 , дијагнозата е определена.

Во случај површината M_1 да сече неколку области B_1, B_2, \dots, B_k можно е да се зборува само за веројатноста на тие S дијагнози.

Соодветната веројатност може да се пресметне како интеграл од густината на веројатноста (x_1, x_2, \dots, x_n) , по површина M_1 , внатре во областите B_1, B_2, \dots, B_k .

$$P(M_1/B_i) = \int_{B_i} p dx_1 dx_2 \dots dx_n, i = 1, 2, \dots, k,$$

каде што $P(M_1/B_i)$ е условна веројатност за M_1 во однос на B_i .

Но, обично интерес претставуваат обратните веројатности, кои ги добиваме на следниот начин.

Да претпоставиме дека површината M_1 сече k -области ($k \leq n$) B_1, B_2, \dots, B_k , односно, при множество од x_1, x_2, \dots, x_n параметри, можни се k -дијагнози.

Користејќи ја формулата на Бајес, добиваме:

$$P(B_i/M_1) = \frac{P(B_i) \cdot P(M_1/B_i)}{\sum_k P(B_k) \cdot P(M_1/B_k)}$$

нека $A_1 = \sum P(B_k) \cdot P(M_1/B_k)$. Тогаш веројатноста априори за дијагнозата B_i ќе биде:

$$P(B_i/M_1) = \frac{1}{A_1} P(B_i) \cdot P(M_1/B_i), i = 1, 2, \dots, k.$$

Со помош на оваа формула можно е да се пресметаат веројатностите на различни дијагнози B_1, B_2, \dots, B_k при постоење на податокот $M_1(x_1, x_2, \dots, x_l)$, а исто така да се определи опаѓачкиот ред чии членови се веројатности на горните дијагнози.

Ако извршиме уште едно испитување и одредиме уште еден параметар x_{l+1} , тогаш добиваме друго геометриско место од точки M_2 , т.е. $M_2(x_1, x_2, \dots, x_l, x_{l+1})$ кое ги сече областите B_1, B_2, \dots, B_k или некоја од нив.

Спрема претходната формула имаме:

$$P(B_i/M_2) = \frac{1}{A_2} P(B_i) \cdot P(M_2/B_i), \text{ каде што}$$

$$A_2 = \sum_k P(B_k) \cdot P(M_2/B_k).$$

Со ова ново испитување и множеството M_2 , може да се одреди друг распоред на дијагнозите од опаѓачки ред или пак претходниот ред да се зајакне.

Од претходно изложеното следува дека неопходно е да се воведат мерка на доверба во дијагнозата.

Да претпоставиме дека параметрите $x_1, x_2, \dots, x_l, \dots, x_n$ се измерени апсолутно точно. Во тој случај можно е да се утврди скоро точно дијагноза и со тоа може да се прекине процесот на одредувањето на M_1, M_2, \dots

Да претпоставиме дека болниот има болест B и нека утврдиме дека веројатноста на тој факт лежи во интервалот

$$1 - a \leq P(B/M_i) \leq 1$$

Тогаш величината $1 - a$ се вика мерка на сигурноста, а величината a се вика мерка на неопределеноста.

Величината a го определува нивото на нашите медицински знаења во даден момент и истовремено го определува обемот на паметењето, во тој смисол колку малку веројатните комбинации е должно да се чуваат. Неточноста во определувањето на параметрите x_1, x_2, \dots, x_l ја намалува веројатноста $P(B/M_i)$ или мерката на сигурност на дијагнозата.

2. Логика на дијагностичкиот процес фундиран на фазниот интервал

Ако точка од фазниот интервал определува состојба на организмот, траекторијата — динамичкиот развој, а определени области од тој простор различни заболувања, нас ќе не интересира што е тоа фазен интервал, врз основа на кој е заснован дијагностичкиот процес.

Нека состојбата на организмот е опишана со симптомите x_1, x_2, \dots, x_n , каде што секој од нив има две значења, и тоа 0 ако го нема симптомот или 1 ако го има.

Тогаш во просторот од симптоми, состојбата на болниот може да се претстави со точката $A(x_1^b, x_2^b, \dots, x_n^b)$, и т.н.

Растојанието помеѓу точките A и B во тој простор можат да се земат по координатите:

$$D_{AB} = \sum_{i=1}^n (x_i^a - x_i^b)$$

$$\text{каде што } |x_i^a - x_i^b| = \begin{cases} 0, & \text{ако } x_i^a = x_i^b \\ 1, & \text{ако } x_i^a \neq x_i^b \end{cases}$$

ако ја посматраме целокупноста на сите симптоми x_1, x_2, \dots, x_n како бинарен број со n бројки, тогаш интервалот помеѓу точките A и B може да се определи со бројот на несовпаднавачки бројки по класите:

$$D_{AB} = \sum_{i=1}^n x_i^a \oplus x_i^b$$

каде што операцијата \oplus е определена со шемата:

\oplus	0	1
0	0	1
1	1	1

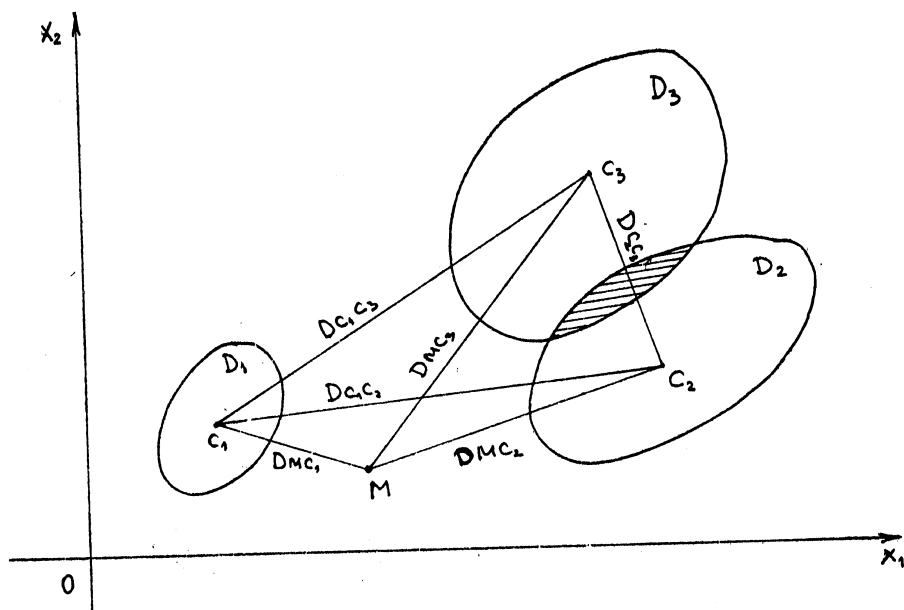
така да, ако m е бројот на несовпаднавачки бројки за интервалот $D_{AB} = m$.

Дефиниција: Дијаметар на множеството D претставува следната вредност:

$$\text{diam } D = \max_{x, y \in D} \{d(x, y)\}$$

Дефиниција: Центарот на множеството е точка која одговара на најтипичниот случај за таа болест.

Така, ја имаме следната геометриска илустрација во R^2 :



На тој начин можно е да се воведо критериум на распознавање на заболувањата.

Проблемот се јавува кога интервалот помеѓу центрите на областите е помал од дијаметарот на областа, т.е. кога областите се сечат можно е да се посматраат како една област.

Критериумот може да се изрази математички со

$$q_{ik} = \frac{DC_iC_k}{\frac{D_i + D_k}{2}}$$

односно, тоа е односот помеѓу растојанијата на центрите на областите и полузбирот од дијаметрите на тие области.

Јасно е дека: $q_{ik} = \begin{cases} < 1 & \text{— областите се сечат} \\ = 1 & \text{— областите се допираат} \\ > 1 & \text{— областите немаат заеднички точки.} \end{cases}$

Постои и група на симптоми која за некои заболувања имаат детерминистички карактер. Некои од нив никогаш не се среќаваат при некои заболувања, а нивното присуство или го исклучува тоа заболување или укажува на присуството на друго заболување.

Како изгледа дијагностичката логика заснована на принципот на фазниот интервал? Да претпоставиме дека при испитувањето е утврдено дека состојбата на болниот се поклопува со точката М.

Дијагнозата се состои во тоа да се определи на која област и припаѓа точката М. Најпрво можеме да утврдиме на кој од центрите на областите точката М им е најблиска. Потребно е да се пресметнат $D_{m_1}, D_{m_2}, \dots, D_{m_k}$ и да се определи најмалата од нив. За да се утврди на која област и припаѓа точката М, се воведува поимот дијагностички радиус — R_j .

Да претпоставиме дека точката М припаѓа на областа Q_j (со центар C_j) и тогаш може да се земе дека М припаѓа на областа Q_j ако $D_{m_j} \leq R_j$.

Дијагностичкиот радиус е аналоген на критериумот T_j во алгоритмот на веројатноста.

Во секоја етапа на дијагностичкиот процес имаме работа со непотполн систем од S симптоми, т.е. со x_1, x_2, \dots, x_s , каде што $s \leq n$, односно со s — димензионален простор, т.е. кај болниот се извршени s испитувања од n можни.

Во врска со тоа при преминување од s — димензионален во $S+1$ — димензионален систем, должината на интервалот се менува, па за споредување на резултатите од една етапа со резултатите од наредната етапа мора да се земат односите меѓу големините, т.е. релативните големини поделени со максималните:

$$d_{AB,S} = \frac{D_{AB,S}}{D_{MAX,S}}$$

каде што $D_{AB,S}$ е растојанието помеѓу А и В во s — димензионален простор, а $D_{MAX,S}$ е максималното растојание во истиот простор. Координатите на центрите не се менуваат, туку само им се додава $s+1$ — вата координата.

Изборот од даденото множество на случаи во една област Q_j може да се оствари на различни начини.

Така на пример, преку минимализацијата на сумата од интервалите на дадените точки до сите други точки од областа

$\sum_{A \in Q_j} = d_{minj} = 1, 2, \dots, s$ (s е број на области) или преку максимализација на производот

$$\prod_{A \in Q_j} (1 - d_{c_j A j}) = d_{max}$$

Бидејќи последниот услов е поосетлив и попрецизен, истиот се користи почесто.

При пресметнувањето на интервалите помеѓу точките или помеѓу дадените точки и центарот кога несовпаднавањата се со иста тежина (симптомите се од исто значење) горните формули даваат задоволувачки резултати.

Во случај кога на секој симптом му е определена неговата тежина, растојанието Dm_{sj} се пресметнува по формулата

$$Dm_{sj} = \sum_{i=1}^n \beta_{ij} \delta_i \quad j = 1, 2, \dots, s,$$

каде што β_{ij} е тежината на симптомите во однос на заболувањата, а

$$\delta_i = \begin{cases} 0 & , \text{ ако симптомите } x_i \text{ се поклопуваат} \\ 1 & , \text{ ако симптомите } x_i \text{ не се поклопуваат.} \end{cases}$$

Вредноста на користењето на оваа формула се состои во тоа што не зависи само од бројот на симптомите, туку и од бројот на областите. Така на пример, додека за една област има детерминистички карактер, за другата неговото учество е минимално.

II. СТАТИСТИЧКИ МОДЕЛ

1. Истражување на клиничкиот преседан

Овој модел се сведува на автоматско истражување на аналогниот или на идентичниот случај, што се јавил порано.

За тоа е потребно да се има добро опремена архива за историјатите на болестите. Потоа се врши споредување на случај со случај од архивата. Оној случај што се поклопува се извлекува. Во случај да нема потполно поклопување, се земаат случаи (или случај) кој е најблизок на разгледуваниот, а со тоа и дијагнозата е определена.

Се воведува поимот на потполно поклопување како однос

$$s = \frac{m}{n}$$

каде што m е бројот на општите, заеднички испитувања за двата случаи (дадениот и оној од архивата), а n е бројот на испитувањата т.е. симптомите земени за дадена класа на заболувања.

Целиот информативен масив (архива) го делиме на два подмасива. Еден од нив ги содржи синдромите, а другиот историјатите на болестите.

На една картичка од ЕСМ е перфорирана првобитната симптоматика, а на друга се наоѓа историјатот на болеста, редниот број на историјатот и други податоци.

Ако машината утврди дека картичките се поклопуваат, дијагнозата е поставена. Ако тоа не е случај, се зема следната картичка со максимално поклопување, т.е. клинички преседан.

При испитувањето на клиничкиот преседан, сите симптоми што се јавуваат во историјатите на болестите, во општ случај се рамноправни, но се можни и случаи каде со помалку карактеристични знаци (симптоми) се определува преседанот.

Во таков случај не се исклучува можноста на појава на дијагностичката грешка. За нејзино избегнување се користи веројатна оценка на симптомите.

Пример: Некоја архива е составена од историјати на болести со заедничка тема „Изненадни промени на крвотокот во мозокот“. Архивот (информативниот масив) се разбива на подкласи по определени дијагнози.

Во секоја подкласа се определува зачестеноста на појавата на i -тиот симптом. Зачестеноста се означува со T_i и ќе биде

$$T_i = \sum_{j=1}^n a_{ji}$$

каде што $a_{ji} = 1$ ако i -тиот симптом се јавува во j -тиот историјат од поткласа со n историјати, а $a_{ji} = 0$ ако не се јавува.

Тежината на симптомот P_i се определува во зависност од зачестеноста на неговата појава и изнесува

$$P_i = \frac{T_i}{n} \quad i = 1, 2, \dots, m,$$

каде што m е бројот на симптомите.

Иако еден ист симптом се јавува често во различни подкласи на дадениот архив, за некои дијагнози од една подкласа тој е покарактеристичен, отколку за други од другата подкласа, а тој наоѓа свој одраз во определувањето на тежината на симптомот.

Во зависност од потполноста на архивот со историјата, тежината на симптомот може да се определи и да биде доста сигурна. На тој начин ЕСМ сама ќе ја оцени тежината на секој симптом и тоа за секоја класа од заболувања.

На крајот како резултат на сето тоа се определува припадноста на состојбата на организмот, т.е. на која подкласа и припаѓа бараната дијагноза. Потоа, во таа класа се определува најверојатен преседан, број на историјатот, клиничка дијагноза на преседанот и други потребни податоци.

Ако архивот не е потполнет, тогаш за зачестеноста и тежината на симптомите можат да бидат запишани (сместени) на магнетен диск и да се чуваат во таква состојба. Така, при определување на дијагнозите не е потребно тие повторно да се пресметнуваат, а со тоа значитно се намалува времето за поставување на дијагнозите.

ЛИТЕРАТУРА

1. Н. С. Мисю к, А. М. Гурлени, В. В. Лозовик: ДИАГНОСТИЧЕСКИЕ АЛГОРИТМЫ, Минск — 1970 год.
2. В. В. Марасанов, Т. Ф. Марасанова, Б. Н. Мягкая МАТЕМАТИЧЕСКИЕ МОДЕЛИ ДИФЕРЕНЦИАЛЬНОЙ ДИАГНОСТИКИ ЗАБОЛЕВАНИЙ, Минск — 1971 год.
3. Група на автори под редакција на академикот АНУССР Н. М. Амосов — МЕДИЦИНСКАЯ ИНФОРМАЦИОННАЯ СИСТЕМА.
4. Д. А. Боровков: КУРС ТЕОРИИ ВЕРОЯТНОСТЕЙ, Москва — 1969.
5. Г. Корн и Т. Корн: СПРАВОЧНИ ПО МАТЕМАТИКЕ.
6. Robert Steven Ledley: USE OF COMPUTERS IN BIOLOGY AND MEDICINE, New York — 1970 год.